



PCT

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/09/701,674A

TIME: 14:28:44

Input Set : A:\pf0539usn_subseqlist.txt

Output Set: N:\CRF4\02112003\I701674A.raw

1 <110> APPLICANT: INCYTE PHARMACEUTICALS, INC.; LAL, Preeti;
 2 YUE, Henry; TANG, Y. Ton;
 3 HILLMAN, Jennifer L.; BANDEMAN, Olga;
 4 CCRLEY, Neil C.; GUEGLER, Karl J.;
 5 GORGONE, Gina A.; BAUGHN, Mariah R.;
 6 PATTERSON, Chandra; LU, Dyung Aina M.
 8 <120> TITLE OF INVENTION: PROTEINS REGULATING GENE EXPRESSION
 10 <130> FILE REFERENCE: PF-0539 USN
 12 <140> CURRENT APPLICATION NUMBER: US 09/701,674A
 C--> 13 <141> CURRENT FILING DATE: 2002-11-14
 15 <150> PRIOR APPLICATION NUMBER: PCT/US99/13281
 16 <151> PRIOR FILING DATE: 1999-06-11
 18 <150> PRIOR APPLICATION NUMBER: US 60/104,624
 19 <151> PRIOR FILING DATE: 1998-10-14
 21 <150> PRIOR APPLICATION NUMBER: US 60/089,079
 22 <151> PRIOR FILING DATE: 1998-06-12
 24 <160> NUMBER OF SEQ ID NOS: 62
 26 <170> SOFTWARE: PERL Program
 30 <210> SEQ ID NO: 1
 31 <211> LENGTH: 379
 32 <212> TYPE: PRT
 33 <213> ORGANISM: Homo sapiens
 35 <220> FEATURE:
 36 <221> NAME/KEY: misc_feature
 37 <223> OTHER INFORMATION: Incyte ID NO: 091290
 39 <400> SEQUENCE: 1
 40 Met Val Phe Phe Thr Cys Asn Ala Cys Gly Gln Ser Val Lys Lys
 41 1 5 10
 42 Ile Ala Val Glu Lys His Val Ser Val Cys Arg Asn Cys Glu Cys
 43 15 20 25 30
 44 Leu Ser Cys Ile Asp Cys Gly Lys Arg Phe Thr His Asp Asn Tyr
 45 35 40 45
 46 Lys Asn His Val Leu Cys Leu Ser His Asp Glu Lys Tyr Gly Gly
 47 50 55 60
 48 Lys Lys Lys Glu Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys
 49 65 70 75
 50 Ala Trp Ile Gln Lys Ile Ser Glu Leu Ile Lys Arg Pro Asn Val
 51 80 85 90
 52 Ser Pro Lys Val Arg Glu Leu Leu His Glu His Ser Ala Phe Asp
 53 95 100 105
 54 Asn Val Pro Arg Lys Lys Ala Lys Phe Gln Asn Trp Met Lys Asn

ENTERED

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/09/701,674A

TIME: 14:28:44

Input Set : A:\pf0539usn_subseqlist.txt

Output Set : N:\CRF4\02112003\I701674A.raw

```

55          110          115          120
56 Ser Leu Lys Val His Asn Glu Ser Ile Leu Asp Gln Val Trp Asn
57          125          130          135
58 Ile Phe Ser Glu Ala Ser Asn Ser Glu Pro Val Asn Lys Glu Gln
59          140          145          150
60 Asp Gln Arg Pro Leu His Pro Val Ala Asn Pro His Ala Glu Ile
61          155          160          165
62 Ser Thr Lys Val Pro Ala Ser Lys Val Lys Asp Ala Val Glu Gln
63          170          175          180
64 Gln Gly Glu Val Lys Lys Asn Lys Arg Gln Lys Lys Glu Glu Arg
65          185          190          195
66 Gln Lys Lys Arg Lys Arg Glu Lys Lys Gln Leu Lys Leu Glu Asn
67          200          205          210
68 His Gln Glu Asn Ser Arg Asn Gln Lys Pro Lys Lys Arg Lys Lys
69          215          220          225
70 Gly Gln Glu Ala Arg Leu Glu Ala Gly Gly Glu Glu Val Pro Glu
71          230          235          240
72 Ala Asn Gly Ser Ala Gly Lys Arg Ser Lys Lys Lys Lys Gln Arg
73          245          250          255
74 Lys Asp Ser Ala Ser Glu Glu Glu Ala Arg Val Gly Ala Gly Lys
75          260          265          270
76 Arg Lys Arg Arg His Ser Glu Val Glu Thr Asp Ser Lys Lys Lys
77          275          280          285
78 Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys Lys
79          290          295          300
80 Asp Glu Ala Pro Ala Lys Gly Lys Phe Asn Thr Lys Gly Thr Lys
81          305          310          315
82 Lys Ala Ile Leu Lys Gln Ala Pro Asp Asn Glu Ile Thr Ile Lys
83          320          325          330
84 Lys Leu Arg Lys Lys Val Leu Ala Gln Tyr Tyr Thr Val Thr Asp
85          335          340          345
86 Glu His His Arg Ser Glu Glu Glu Leu Leu Val Ile Phe Asn Lys
87          350          355          360
88 Lys Ile Ser Lys Asn Pro Thr Phe Lys Leu Leu Lys Asp Lys Val
89          365          370          375
90 Lys Leu Val Lys
91 <210> SEQ ID NO: 1
92 <211> LENGTH: 130
93 <212> TYPE: PRT
94 <213> FEATURE:
95 <214> NAME/KEY: misc feature
96 <223> OTHER INFORMATION: Incyte ID No: 815856
97 <400> SEQUENCE:
98 Met Phe Gly Thr Pro Gln Glu His Arg Asn Met Pro Gln Ala Asp
99 Ala Met Val Leu Val Ala Arg Asn Tyr Gln Arg Lys Lys Asn Glu
100 Cys Arg Gln Lys Gln Arg Gln Glu Ile Ala Arg Gln Ala Ala Lys

```

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/09/701,674A

TIME: 14:28:44

Input Set : A:\pf0539usn_subseqlist.txt

Output Set: N:\CRF4\02112003\I701674A.raw

```

109      35      40      45
109 Met Ala Asp Glu Ala Ile Leu Gln Glu Arg Gln Arg Gly Gly Pro
110      50      55      60
111 Glu Glu Gly Val Arg Gly Gly His Pro Pro Ala Ile Gln Ser Leu
112      65      70      75
113 Ile Asn Leu Leu Ala Asp Asn Arg Tyr Leu Thr Ala Glu Glu Thr
114      80      85      90
115 Asp Lys Ile Ile Asn Tyr Leu Arg Glu Arg Lys Glu Arg Leu Met
116      95     100     105
117 Arg Ser Ser Thr Asp Ser Leu Pro Gly Glu Leu Arg Gly Arg Ala
118     110     115     120
119 Glu Ala Arg Phe Pro Ala Asn His Ser Gly Arg Pro Arg Val Pro
120     125     130     135
121 Arg
122 <210> SEQ ID NO: 3
123 <211> LENGTH: 230
124 <212> TYPE: PRT
125 <213> ORGANISM: Homo sapiens
126 <220> FEATURE:
127 <221> NAME/KEY: misc_feature
128 <223> OTHER INFORMATION: Incyte ID No: 996352
129 <400> SEQUENCE: 3
130 Met Ser Ser Ser Tyr Tyr Val Asn Ala Leu Phe Ser Lys Tyr Thr
131      1      6
132 Ala Gly Ala Val Leu Ile Gln Asn Ala Glu Ile Thr Ser Tyr Ser
133      11     16
134 Phe Ala Pro Asn Ser Gln Arg Ser Gly Tyr Gly Ala Gly Ala Gly
135      21     26
136 Ala Phe Ala Ser Thr Val Pro Gly Leu Tyr Asn Val Asn Ser Pro
137      31     36
138 Leu Tyr Gln Ser Pro Phe Ala Ser Gly Tyr Gly Leu Gly Ala Asp
139      41     46
140 Ala Tyr Gly Asn Leu Pro Cys Ala Ser Tyr Asp Gln Asn Ile Pro
141      51     56
142 Gly Leu Cys Ser Asp Leu Ala Lys Gly Ala Cys Asp Lys Thr Asp
143      61     66
144 Gln Gly Ala Leu His Gly Ala Ala Glu Ala Asn Phe Arg Ile Tyr
145      71     76
146 Arg Trp Met Arg Ser Ser Gly Pro Asp Arg Lys Arg Gly Arg Gln
147      81     86
148 Thr Tyr Thr Arg Tyr Gln Thr Leu Gln Leu Gln Lys Glu Phe His
149      91     96
150 Phe Asn Arg Tyr Leu Ile Arg Arg Arg Arg Ile Glu Ile Ala His
151     101    106
152 Ala Leu Cys Leu Thr Glu Arg Gln Ile Lys Ile Trp Ile Gln Asn
153     111    116
154 Arg Arg Met Lys Trp Lys Lys Glu His Lys Asp Glu Gly Pro Trp
155     121    126
156 Ala Ala Ala Ala Pro Gln Gly Ala Val Pro Ser Ala Ala Ala Thr

```

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/09/701,674A

TIME: 14:28:44

Input Set : A:\pf0539usn_subseqlist.txt

Output Set: N:\CRF4\02112003\I701674A.raw

```

161          200          205          210
162 Ala Ala Ala Asp Lys Ala Asp Glu Glu Asp Asp Asp Glu Glu Glu
163          215          220          225
164 Glu Asp Glu Glu Glu
165          230
166 <210> SEQ ID NO: 4
167 <211> LENGTH: 131
168 <212> TYPE: PRT
169 <213> ORGANISM: Homo sapiens
170 <220> FEATURE:
171 <221> NAME/KEY: misc feature
172 <223> OTHER INFORMATION: Incyte ID No: 1273778
173 <400> SEQUENCE: 4
174 Met Ser Glr Val Thr Phe Ser Asp Val Ala Ile Asp Phe Ser His
175      1          5          10          15
176 Glu Glu Trp Ala Cys Leu Asp Ser Ala Gln Arg Asp Leu Tyr Lys
177      20          25          30
178 Asp Val Met Val Glr Asn Tyr Glu Asn Leu Val Ser Val Gly Leu
179      35          40          45
180 Ser Val Thr Lys Pro Tyr Val Ile Met Leu Leu Glu Asp Gly Lys
181      50          55          60
182 Glu Pro Trp Met Met Glu Lys Lys Leu Ser Lys Ala Tyr Pro Phe
183      65          70          75
184 Ser Leu Ser Ser Ser Val Pro Asn Leu Val Asn Ile Tyr Ile Thr
185      80          85          90
186 Ala Leu Phe Glu His Cys Ser Glu Val Thr Glu Ile Phe Glu Leu
187      95          100          105
188 Ser Glu Leu Cys Val Phe Trp Val Leu His Phe Leu Ser Asn Ser
189      110          115          120
190 Pro Asn Ser Thr Val Glu Ala Phe Phe Lys Lys
191      125          130
192 <100> SEQ ID NO: 3
193 <211> LENGTH: 411
194 <212> TYPE: PRT
195 <213> ORGANISM: Homo sapiens
196 <220> FEATURE:
197 <221> NAME/KEY: misc feature
198 <223> OTHER INFORMATION: Incyte ID No: 1509715
199 <400> SEQUENCE: 4
200 Met Ser Lys Arg Pro Ser Tyr Ala Ile Lys Ile Thr Ile Ala Pro
201      1          5          10          15
202 Ala Thr Ala Met Pro Ser Thr Pro Gly Pro Val Gly Tyr Asn Pro
203      20          25          30
204 Tyr Ser His Leu Ala Tyr Asn Asn Tyr Arg Ile Gly Gly Asn Pro
205      35          40          45
206 Glu Thr Asn Ile Arg Ile Asn Ile Ser Ile Ile Ile Ile Ile Ile
207      50          55          60
208 Lys Pro Pro Lys Pro Pro Asp Lys Pro Leu Met Pro Tyr Met Arg
209      65          70          75

```

RAW SEQUENCE LISTING

DATE: 02/11/2003

PATENT APPLICATION: US/09/701,674A

TIME: 14:28:44

Input Set : A:\pf0539usn_subseqlist.txt

Output Set: N:\CRF4\02112003\I701674A.raw

```

216 Tyr Ser Arg Lys Val Trp Asp Gln Val Lys Ala Ser Asn Pro Asp
217                               80                               85                               90
218 Leu Lys Leu Trp Glu Ile Gly Lys Ile Ile Gly Gly Met Trp Arg
219                               95                               100                               105
220 Asp Leu Thr Asp Glu Glu Lys Gln Glu Tyr Leu Asn Glu Tyr Glu
221                               110                               115                               120
222 Ala Glu Lys Ile Glu Tyr Asn Glu Ser Met Lys Ala Tyr His Asn
223                               125                               130                               135
224 Ser Pro Ala Tyr Leu Ala Tyr Ile Asn Ala Lys Ser Arg Ala Glu
225                               140                               145                               150
226 Ala Ala Leu Glu Glu Glu Ser Arg Gln Arg Gln Ser Arg Met Glu
227                               155                               160                               165
228 Lys Gly Glu Pro Tyr Met Ser Ile Gln Pro Ala Glu Asp Pro Asp
229                               170                               175                               180
230 Asp Tyr Asp Asp Gly Phe Ser Met Lys His Thr Ala Thr Ala Arg
231                               185                               190                               195
232 Phe Gln Arg Asn His Arg Leu Ile Ser Glu Ile Leu Ser Glu Ser
233                               200                               205                               210
234 Val Val Pro Asp Val Arg Ser Val Val Thr Thr Ala Arg Met Gln
235                               215                               220                               225
236 Val Leu Lys Arg Gln Val Gln Ser Leu Met Val His Gln Arg Lys
237                               230                               235                               240
238 Val Glu Ala Gln Ser Leu Glu Ile Glu Glu Arg His Gln Glu Lys
239                               245                               250                               255
240 Lys Arg Lys Phe Ser Glu Ser Thr Asp Ser Phe Asn Asn Glu Ile
241                               260                               265                               270
242 Lys Arg Leu Cys Gly Leu Lys Val Glu Val Asp Met Glu Lys Ile
243                               275                               280                               285
244 Ala Ala Glu Ile Ala Gln Ala Glu Glu Gln Ala Arg Lys Arg Gln
245                               290                               295                               300
246 Glu Glu Arg Glu Lys Glu Ala Ala Glu Gln Ala Glu Arg Ser Gln
247                               305                               310                               315
248 Ser Ser Ile Val Pro Glu Glu Glu Gln Ala Ala Asn Lys Gly Glu
249                               320                               325                               330
250 Gln Lys Lys Asp Asp Glu Asn Ile Pro Met Glu Thr Glu Glu Thr
251                               335                               340                               345
252 His Leu Glu Gln Thr Thr Glu Ser Gln Glu Asn Gly Glu Glu Gly
253                               350                               355                               360
254 Thr Ser Thr Phe Glu Asp Lys Glu Ser Gly Gln Glu Gly Val Asp
255                               365                               370                               375
256 Ser Met Ala Glu Glu Gly Thr Ser Asp Ser Asn Thr Gly Ser Glu
257                               380                               385                               390
258 Ser Asn Ser Ala Thr Val Glu Glu Pro Phe Thr Asp Pro Ile Pro
259                               395                               400                               405
260 Ser Asp Thr Lys Lys Glu
261                               410
262 <210> SEQ ID NO: 6
263 <211> LENGTH: 226
264 <212> TYPE: PRI

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/09/701,674A

DATE: 02/11/2003
TIME: 14:28:45

Input Set : A:\pf0539usn_subseqlist.txt
Output Set : N:\CRF4\02112003\I701674A.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:6; Xaa Pos. 218,223
Seq#:35; N Pos. 6,61,882,883,884,885,886,887,888,889,890,891,892,893,894
Seq#:35; N Pos. 895,896,897,898,899,900,901,902,903,904,905,906,907,908,912
Seq#:35; N Pos. 914,916,920,922,929,932,964,967,971,975,985,1000,1028,1033
Seq#:35; N Pos. 1039,1052,1053,1063,1065,1071,1080,1090,1100,1102,1110,1113
Seq#:35; N Pos. 1123,1128,1134,1139,1142,1143,1144,1153,1158,1164,1169,1177
Seq#:35; N Pos. 1180,1181,1182,1184,1189,1191,1192,1211,1224,1239,1246,1247
Seq#:35; N Pos. 1243,1263,1266,1274,1284,1291,1314,1321,1322,1339,1341,1343
Seq#:35; N Pos. 1347,1351,1353,1359,1362,1363,1364
Seq#:37; N Pos. 684,693,740,755,757,776,796,805,811,824,850,856
Seq#:42; N Pos. 2223
Seq#:45; N Pos. 61,63,69,92,100,151,165,179,182,198,218
Seq#:51; N Pos. 1957
Seq#:59; N Pos. 1203,1204